DOCUMENT RESUME

ED 383 185 FL 022 971

AUTHOR Laurier, Michel D.; Chung How, Mee Lian
TITLE The Use of Verbal Protocols To Validate an Adaptive Test.
PUB DATE Mar 95
NOTE 15p.; Paper presented at the Annual Language Testing Research Colloquium (Long Beach, CA, March 1995).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Adaptive Testing; Advanced Placement; *Computer Assisted Testing; Foreign Countries; *French; Higher Education; *Language Tests; *Oral Language; Screening Tests; Student Evaluation; Testing; *Test Validity

ABSTRACT

This paper presents the validation of an adaptive test developed for placement purposes in French at the post-secondary level, the French CAPT. Results from a second experiment are presented in which verbal protocols were obtained from nine college students of various levels with the thinking-aloud technique used during completion of the exam. The 4-part test includes a paragraph reading section with a comprehensive question about each paragraph, a situation reading section with grammatical statement selection, a sentence gap section, and a semi-authentic dialogue with questions section. A 5th part, under development, will use a graded-response model for self-assessment of ability. All questions are multiple choice; items in the first three parts have been calibrated with a IRT 3-parameter model. Administration of the French CAPT is similar to other computerized adaptive tests and uses a stradaptive algorithm, programmed for the selection of items. Findings indicate that perception of difficulty varied considerably, specifically regarding subjective vs objective difficulty. It is concluded that this use of verbal protocols evidenced no effect due to presentation mode and the use of appropriate desired mental processes during the test. (Contains 27 references.) (NAV)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Michel D. Laurier
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Language Testing Research Colloquium
Long Beach, CA, March 1995

The Use of Verbal Protocols to Validate an Adaptive Test

Michel D. LAURIER & Mee Lian CHUNG HOW
University of Montreal

ED 383 185

The most important contribution of cognitive psychology is probably the emphasis on information processing. Learning is not simply defined in terms of outcome but also in term of processes. As shown by Snow and Lohman (1993) the focus on processes has important implications on test design. As far as multiple choice items are concerned, it is now clear that simple response analysis has to be complemented with information about the way these responses are generated (Mislevy, 1993). Many testing theorists believe that thinking processes should play a role in the representation of test constructs and that verbal protocols are very helpful in the description of the thinking processes (Cronbach, 1971; Embretson, Schneider & Roth, 1986; Messick, 1989).

Although the value of verbal protocols is recognized, they have been used in only a few studies. Bloom & Broder (1950) and Kropp (1956) were among the first researchers to use verbal protocols to understand the reasoning processes that are involved in test tasks. These studies were then followed in the sixties by researches made by McGuire (1963) and Connolly & Wantman (1964) who both attempted to compare subjects' reasoning with expert reasoning.

With the work of Ericsson and Simon (1984, 1987), the elicitation and analysis techniques became more refined and more systematic. Norris (1992) describes a method for using verbal protocols to demonstrate how critical thinking can occur in multiple choice tests. For Norris, the technique is integrated in the item analysis procedure which aims at writing better test items.

FL022971

In second language testing, verbal protocols collected from examinees have also been used. Cohen (1984) reports some studies that have been done on the strategies used by examinees on language tests: guessing, translation, word matching... He mentions that the type of strategies is related to the students' level and the nature of the task. Grotjahn (1986, 1987) favors a combination of psychometric analysis and qualitative analysis in studying test-taking processes. He recommands the thinking-aloud technique and the retrospective interview. He also illustrates the value of this approach as a complement to the validation of a C-test (Klein-Bradley, 1985). Feldman & Stemmer (1987) also examine the C-test by means of thinking-aloud and retrospective data. They establish a list of strategies used on a C-test and propose a tentative model of the problem-solving process involved in this type of test.

BEST COPY AVAILABLE

THE ADAPTIVE TEST

The present study is part of the validation of an adaptive test that we have built for placement purposes in French, at the post-secondary level.[1] A prototype of this test, the *French CAPT*, is operational. The test presently includes four parts:

1) The student reads short paragraphs (about 30 words) and answers a comprehension question about each paragraph.
2) The student reads a situation (in L1) and selects, among four grammatically correct statements (in L2), the most appropriate one.
3) The student fills the gap in different sentences (vocabulary and grammar items).
4) The student hears semi-authentic dialogues (about 2 minutes) and answers 3 questions on each dialogue.

All the questions are multiple choice and the items are stored in four different banks - one for each part. The items in the first three parts have been calibrated with a IRT three-parameter model (Birnbaum, 1968). For the first three parts, the model fits fairly well in spite of some departures of the unidimensionlity assumption on the second and third parts (Blais & Laurier, in press). However, this assumption could not be met with the fourth part because of dimensionality problems related to the clustering of items around dialogues. A testlet approach (Wainer & Kiely, 1987) has then been used and each dialogue has been calibrated using a two-parameter graded-response model (Samejima, 1978). A fifth part is under development. It will also use the graded-response model as students will be asked to refer to a rating scale to self-assess their speaking ability in different situations.

The administration of the *French CAPT* is similar to the administration of other computerized adaptive tests (CAT) that have been developed in second language (Larson & Madsen, 1985). A stradaptive algorithm has been programmed for the selection of the items (Vale & Weiss, 1974). The items are selected so that they are neither too difficult nor too easy and the ability estimation is revised after each answer. The procedure goes on until the error of measurement is acceptable or the maximum number of items is reached. As a result, the *French CAPT* is shorter than a conventional test. Prior information from student's background or from the preceding parts is used to select the first item at the beginning of each part. Because of the grouping of the items on the basis of the task type, the results are reported as a profile and as a general proficiency level with respect of 14 possible levels ranging from *Absolute Beginner* to *Very Advanced +*.

This study follows a first one that has been conducted on the examinees' reactions towards the *French CAPT* (Laurier, 1993). Prior to this first study, comparaisons had been made between the first three parts of the *French CAPT* and

a conventional paper-and-pencil version constructed with the same item banks. Correlations were high in spite of some placement differences that seemed to be caused by the scoring procedure - number of right answers on the conventional version VS maximum likelihood on the CAT version. In our first study, we wanted to kwow if there was any evidence of a method effect related to the administration mode that could be found in students' reactions towards the test. Therefore, a questionnaire and a retrospective discussion have been used to analyse students' perceptions on both versions. Surprisingly, the analysis did not show any major differences in students' perceptions on aspects such as difficulty, duration or test anxiety. However this first study suggested that the test strategies that are currently used on language tests do not work in the same way. As a conclusion to this first study, we realized that it should be complemented with a verbal protocol analysis using the thinking-aloud technique. An additional study was even more necessary since this one was restricted to the first three parts and, therefore, there was no information about the processes which undergo during the CAT administration of a listening test.

This paper reports the results of the second experiment in which verbal protocols were obtained from students of various levels with the thinking-aloud technique while they were doing the *French CAPT*. The purpose is to highlight the strategies that are used by the students at different levels on a language placement CAT. More specifically, it aims at determining if there is any mode effect (as described by Steinberg, Thissen & Wainer, 1990) that could influence the mental processes and affect the validity of the test.

METHODOLOGY

This experiment has been carried out with 9 subjects, enrolled in a French second language programme at the University of Montreal. The nine students were of different levels, ranging from *Beginner* to *Very Advanced*. Even if all the nine students had a good knowledge of English, for most of them, it was a second language.

Each student has been asked to do the *French CAPT*, thinking aloud to describe how they get to an answer or interpret the input. They were allowed to use either English or French. The observer explained in detail what the subject was expected to do. In addition to the directions, Ericsson & Simon (1987) recommend a training session for the subjects. However, because of the nature of the task, it was not possible to allow time for practice. The observer tried to be as discreet as possible, keeping in mind the information that had to be collected. While doing the test, students were asked to comment on the difficulty of the task, the different strategies they used to answer the questions, and their comprehension of the input. At the end

of the test, they were asked to comment on the result. The test was done by each subject individually with no time limit. The students spent about one hour and forty-five minutes doing the test.

The entire session was recorded on a VHS videotape. The video input was used to record a converted VGA signal coming from the computer screen; the audio input was used to record the observer's interventions, the dialogues coming from the sound board through the speakers and, of course, the students' comments.

The transcripts were analysed to point out the recurring comments made by the students in doing this computerized test. Following the suggestion of Miles & Huberman (1994), we classified the recorded data according to the most relevant content categories.

RESULTS

The results of the classification are summarized in a table in the Appendix. Eight major categories were established.

Difficulty : This was quite surprising. The advanced students found the test too easy whereas, for the beginners, it was too difficult. As it is an adaptive test, this should not happen for we expect students to find it neither difficult nor easy. Moreover, the beginners who did the test, skipped several questions which they found too difficult to solve. Each student was supposed to answer questions suited to their actual proficiency level in French and it seems that the students, especially beginners, didn't feel that way at all.

I.L.: "*Can I change the level, or should I pass the level ? because this level is too difficult for me.*"

It also appears that the test, which consists of four different parts, presents more difficulties in certain parts. For most of the students, Part I was the most difficult because of the complex vocabulary, and the multiple-choice answers were not too obvious.

M.M.: " *Part I, more difficult because if you can't understand the words, you can't understand the paragraph.*"

Even for some advanced students, they were rather ambiguous.

F.R.: " *les réponses se ressemblaient et j'étais pas trop sûre.*"

The easiest part seemed to be the listening test (Part IV) and neither beginners nor advanced students had great trouble with this part except for one student who had problems with doing two things at the same time, namely listening and reading. Another student also pointed out that there was no time to think because one must listen to the dialogue, keep in mind the answers and answer straight away before forgetting the dialogue.

C.L.: " *J'ai trouvé la dernière partie la plus difficile, parce que là, tu peux pas réfléchir... Il faut écouter une seule fois et puis répondre.*"

Accuracy : For most of the students, the level to which they were assigned to, at the end of the test was what they expected. For beginners and advanced students, there were no surprise at all. However, some intermediate students were more surprised by the individual results in the different parts than by their overall results. For example, this student knew that she didn't score very well in the test but was greatly surprised when she saw that in the listening test (Part IV) she had a perfect score.

D.B.: " *(Part I) not surprising because vocabulary way overhead of me. (Part III) my grammar, oh.... that doesn't surprise me. That (Part IV) surprises me. I wouldn't expect to... I didn't think I do that well.*"

Specific strategies : We were interested in finding out the different strategies used by the students in the four different parts when they didn't know an answer.

Part I : The students refered to the general context of the paragraph to get a general idea, even if there were words unknown to them. Thus, they tried to find a link between the paragraph and one of the options given. Very often, they also made use of the words they knew and tried to compare them in the answers.

D.B.: " *I am taking the words I do know and I, kind of looking at the words I don't know and looking at it in the context that it's in, and try with the words I know to make something up...*"

Sometimes, students also chose by pure intuition if they could not figure out the right answer at all.

N.H.: " *This question, kind of, bothers me, because I don't know what it is, at all, at all, ... I'm just going to go at "d" because it's the shortest, the simplest and the sweetest.*"

Beginners skipped questions which contained too many complex words, unknown to them, because they couldn't understand anything and they felt it was of no use to try the answers as they didn't understand the general idea of the text.

Part II : There was not much problem with this part as the text was in English and all students understood most of it. The text presents mainly socio-linguistic situations where they had to use their personal experience. They refered to what they use to hear from their familiar environment, or to what they use to say if they are in similar situations.

J.J.: " *et comme ça, je cherche dans ma tête, les phrases que j'ai entendues le plus souvent.*"

Part III : For this part, the task was to find the missing word in a sentence. Most students who didn't know the right word, chose one by the sound. If it sounded right in the sentence, they just took it even if they didn't know the meaning of the word.

D.B.:"*I am putting each word in. O.K. I did tell you my grammar's poor, right. So, I'm trying to listen to it, some things sound right to me, b, it sounded good in my head.*"

N.H.: " *So I felt I was guessing in part III but perhaps I got it right because I'm used to hearing it. I say it loud, to hear what it sounds like.*"

Part IV : The listening test didn't present great difficulties for the students, even if the dialogues were heard only once. Students first read the questions and the options, then listened to the dialogue and at the end of the dialogue, answered the questions. While listening to the dialogue, they tried to discard the most unlikely distractors. For beginners, they listened mainly to words and tried to find them in the answers. However, for students of all levels, if they missed an answer and didn't hear it while listening, they used logic and tried to figure out what was right according to what they had heard in general.

M.M.: " *Now you see, I don't remember. What I can do, I can think. ... I think it's about helping him with his work because that makes more sense.*"

Length & Computerized Environment: The overall impression was that the computerized test was not long and some even found it quite short. As for the computerized environment, it was surprising to find that some advanced students had some problems with the keyboard, whereas the rest of the subjects seemed to be quite at ease with a computer. Some even said that they did not feel any kind of stress and were quite relaxed while doing the test. Only one advanced student pointed out that she would have preferred a conventional paper-and-pencil test to a computerized one.

C.L.: " *Je peux pas penser. Je trouve que c'est difficile quand c'est sur ordinateur. Je me sens plus à l'aise si c'est sur une feuille.*"

Backtracking : Another important element was to find out if students go back to the text or answers when they are doing a test. All students had a tendency to go back to reading again the text when it was not clear at the first reading, or to check an answer. Many students mentionned that it was not possible to review their answers once they were given. However, few students were interested in going over their answers at the end of the test.

Test strategies : One common element was present in all the examinees' recorded data. All students, regardless of their level, used their mother-tongue to help them to answer questions. Translation in the student's mother-tongue was a recurring strategy and even for some students, grammatical rules in their mother-tongue were used as a support because they were similar to those in French. Students' mother-tongue seems to be an important source on which second language learners rely; they feel that it is of great help to them.

C.L.: " *Normalement, je pense en suédois, c'est la même chose.*"

M.M.: " *I try to translate and according to that, compare with the text in English. This is easier because I have access too English, I can understand 100% the meaning of the paragraph.*"

I.L.: " *I translate in my native language, to Russian, I translate to*

Russian...yes. It's too difficult for me to think in French, so I translate in Russian."

Another strategy was to use their own knowledge of French grammar or oral French, that is words or idiomatic phrases that they heard before or used to hear quite often in conversation and which sounded good to them. This is quite effective for students succeed in finding the right answer.

Instructions : For this test, all instructions are given on computer and we were interested in finding out if students had problems with them. It seems that the instructions are clear and direct and students didn't encounter difficulties while reading them. However, we have to point out that all students read attentively the instructions at the beginning of the four main parts, but afterwards, they forgot about the instructions given before each item for it was just a repetition.

N.K.: " *I'm skipping the red now. I figure it's the instructions, it says the same thing every time."*

F.R.: " *Oh, non! je les lis jamais. Je les lis pas car il dit toujours la même chose."*

DISCUSSION

An interesting aspect that is observed is the perception of the difficulty. The students' comments illustrate the importance in CAT of the distinction between a subjective difficulty and objective difficulty (Prestwood & Weiss, 1977). Subjective difficulty is related to the cognitive charge of the task and the importance of perceived weaknesses in doing the task whereas objective difficulty is a statistical result from the calibration procedure.

This distinction explains why students generally say that Part I (paragraph reading) is perceived as being much more difficult than Part IV (listening comprehension) even though the comprehension questions are not objectively more difficult. It seems that Part I is more cognitively demanding. One major factor that seems to contribute to the subjective difficulty is the vocabulary. The complexity of the task is related to the knowledge of lexical elements. Beginners who have a very limited vocabulary tend to believe that a task is more difficult than what the objective difficulty index may suggest when they are faced with new words.

In conjunction with the distinction between objective and subjective difficulty, we were surprised to find that students do not seem to realize the adaptiveness of the test. There were almost no comment regarding the selection of the items according to the level of ability. This suggests that the psychological advantage of adaptive testing that prevents frustration because of items that are supposed to be neither too difficult nor too easy has been overemphasized.

The analysis of the protocols confirmed the finding in our first study about

guessing. Even at the beginner's level, students do not guess whenever they are not sure about the correct answer. A typical examinee eliminates the most unlikely options and always finds an indication leading to the answer he/she believes to be the right one. Translation is also a common strategy. Even though one might expect only beginners to use this strategy, advanced students as well as beginners use translation whenever they are not totally sure about the answer. Lord (1980) was certainly right when he labelled the 3rd parameter "pseudo-guessing". Although the use of this parameter with multiple-choice questions improves the model fit, it is not clear at all what this parameter really represents.

Verbal protocols are also a valuable source of information in determining the mental processes that are actually going on during the execution of a test task. We were particularly interested in knowing what were the processes put into play during the listening comprehension part. Since listening abilities are difficult to observe, one may believe that well-formulated multiple-choice questions provide a good measure of comprehension. Distractors should correspond to real hypotheses on a passage interpretation. Therefore, a good test task is a problem-solving procedure that consists in determining the most probable interpretation. Verbal protocols analysis should provide evidence that this procedure is actually taking place. As far as Part IV of *French CAPT* is concerned, we found that the students first read the options and then used to discard the distractors while listening to the dialogue. As predicted, retrieving simple information seems to be a simpler task than inferencing.

We were mainly interested in finding if there was a method effect related to the presentation mode. The verbal protocols in this study prove that there was no major effect due to the computerized administration. Students were not more anxious because of the computer. Only two had minor problems with the keyboard. The directions were generally well understood. The only comment that could indicate some effect of the presentation mode regards the possibility of going back to the previous questions in order to use clues for other questions. Some students have mentioned that this was not possible but said that they would not have used this possibility anyway. In fact, this is an asset from a psychometric point of view because it ensures that the local independence assumption is met.

In conclusion, the use of verbal protocols has been very useful in the validation of the *French CAPT*. It has shown that there is no effect due to the presentation mode and that the mental processes going on were those that we were interested in. One could certainly wonder if the thinking aloud technique changes the nature of the task. Even though Norris (1990) has shown that this is not the case for a test of critical thinking based on multiple-choice questions, one could suspect that verbalization is a heuristic strategy that facilitates the response to language tasks. This could be the reason why some intermediate students were placed at a higher level than what they had expected: thinking-aloud helped them to find the right answer. In spite of this limitation, we believe that verbal protocols should be included in the validation of any language test.

REFERENCES

Bachman, L.F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In Lord, F.M. & Novick, M.R., *Statistical Theories of Mental Test Scores. Reading*, MA: Addison-Wesley.

Blais, J.G. & Laurier, M.D. (In press). The dimensionality of a language placement test from several analytical perspectives. *Language Testing 12*, no 1.

Bloom, B.S. & Broder, J.L. (1950). *Problem-solving processes of college students*. Chicago: University of Chicago Press.

Cohen, A.D. (1984). On taking tests: what the students report. *Language Testing,1*,1, 70-81.

Cronbach, L.J. (1971). Test validation. In Thorndike, R.L. (Ed.), *Educational measurement* (2nd ed.) Washington, D.C: American Council on Education. 443-507.

Embretson, S., Schneider, L.M.,& Roth, D.L.(1986). Multiple processing strategies and the construct validity of verbal reasoning tasks. *Journal of Educational Measurement, 23*, 13-32.

Ericsson, K.A. & Simon, H.A. (1984). *Protocol analysis: verbal reports as data*. Cambridge, MA: MIT Press.

Ericsson, K.A. & Simon, H.A. (1987). Verbal reports on thinking. In Faerch, C. & Kasper, G. (Ed.) *Introspection in Second Language Research*. Clevedon, Philadelphia: Multilingual Matters Ltd. 24-53.

Feldmann, U. & Stemmer, B. (1987). Thin__aloud a__retrospective da__in C-te__ taking: diffe__ languages - diff__ learners - sa__ approaches? In Faerch, C. & Kasper, G. (Ed.) *Introspection in Second Language Research*. Clevedon, Philadelphia: Multilingual Matters Ltd. 251-66.

Grotjahn, R. (1986). Test validation and cognitive psychology: some methodological considerations. *Language Testing, 3*,2, 159-85.

Grotjahn, R. (1987). On the methodological basis of introspective methods. In Faerch, C. & Kasper, G. (Ed.) *Introspection in Second Language Research*. Clevedon, Philadelphia: Multilingual Matters Ltd. 54-81.

Huberman, A.M. & Miles, M.B. (1994). *Qualitative data analysis: an expanded sourcebook*. (2ed.) Thousand Oaks, London: Sage Publications.

Klein-Bradley, C. (1985). A cloze-up on the C-test: a study in the construct validation of authentic tests. *Language Testing, 2*, 1, 76-104.

Larson, J.W. & Madsen, H.S. (1985). Computerized adaptive testing: Moving beyond Computer-Assisted Testing. *CALICO Journal*, March 1995, 32-43.

Laurier M.D. (1993) Un test adaptatif en langue seconde: la perception des apprenants. In Hivon, R. (Ed.) *L'évaluation des apprentissages: réflexions, nouvelles tendances et formation*. Sherbrooke, Canada: Éditions du CRP, 97-118.

Lord, F.M. (1980). *Applications of Item Response Theory to Practical Testing Problems*. Hillsdale, NJ: Lawrence Erlbaum.

Messick, S.(1989). Validity. In Linn, R.L. (Ed.), *Educational measurement* (3rd ed.) New York: Macmillan. 13-103.

Mislevy, R.J. (1993). A framework for studying differences between multiple choice and free response test items. In Bennett, R.E. & Ward W.C. (Eds), *Construction versus choice in cognitive measurement: Issues in constructed responses, performance testing, and portfolio assessment*. Hillsdale, NJ: Lawrence Erlbaum, 54-87.

Norris, S.P. (1990). Effect of elicting verbal reports of thinking on critical thinking test performance. *Journal of Educational Measurement, 27*,41-58.

Norris, S.P. (1992). A demonstration of the use of verbal reports of thinking in multiple-choice critical thinking test design. *The Albertan Journal of Educational Research*, 38, 3, 155-176.

Prestwood, J.S. & Weiss, D.J. (1977). *Accuracy of perceived test-item difficulties*, (R.R. 77-3). Minneapolis, MN: University of Minnesota, Dept. of Psychology.

Samejima, F. (1978). The application of graded-response models: The promise for the future. In Weiss, D.J. (Ed.) *Proceedings of the 1977 Conference on Computerized Adaptive Testing*. Minneapolis, MN: University of Minnesota, 28-37.

Snow, R.E. & Lohman, D.F. (1993). Cognitive psychology, new test design, and new test theory: An introduction. In Fredericksen, R.L., Millevy, R.J & Bejar I.I. (Eds), *Test theory for a new generation of tests*. Hilldale, NJ: Lawrence Erlbaum, 127-176.

Steinberg, L. Thissen D. & Wainer, H. (1990). Validity. In Wainer, H. (Ed). *Computerized adaptive testing: A primer*. Hillsdale, NJ: Lawrence Earlbaum.

Vale, C.D. & Weiss, D.J. (1975). *A study of computer-administered stradaptive ability testing*, (R.R. 75-4). Minneapolis, MN: University of Minnesota, Dept. of Psychology.

Wainer, H. & Kiely, G.L. (1987). Item clusters and computerized adaptive testing: A case for testlets. *Journal of Educational Measurement, 24*, 185-201.

1. The development and validation of the French CAPT has received the financial support of the Social Sciences and Humanities Research Council of Canada. (SSHRC).

SUBJECTS	DIFFICULTY	ACCURACY	SPECIFIC STRATEGIES	LENGTH	COMPUTE. ENVIRON.	TEST STRATEGIES	BACK-TRACKING	INSTRUC-TIONS
S.A. (Beginner)	Difficult	As expected. *Actually I think I didn't know much.*	General context. Known words. Skipped several items.	Short *It's faster this way.*	At ease. *I don't feel nervous here.*	Translation and key words. *I try to translate words I know.*	Document	NC
I.L. (Near-beginner)	Not his level. *This level is too difficult for me.*	As expected. *I expect such results. Good results for me.*	General context. Skipped some items. *This..... is very difficult and I skip the question.*	NC	NC	Translation in L1. *It's too difficult for me to think in French, so I translate to Russian.*	Document	Clear
D.B. (Low Intermed.)	Part IV easier. *...listening easier, I listen to French a lot.*	As expected, except Part IV. *That surprises me. I didn't think I do that well.*	Part I : Context Part II :Intuition and personal experience. Part III :- By the sound. *....some things do sound right to me, b, it sounded good in my head.*	NC	Tired. *Your eyes get tired when you sit so close to the screen.*	Translation in L1. Personal experience. *Translation in my head as the test was going. Making my little diagram in my head.*	Document *I keep refering back to what I'm doing, that is to the paragraph.* Réponses. *You can't go back and do it over again.*	Clear
M.M. (High Intermed.)	Part I more difficult. *Part I, more difficult, because if you can't understand words, you can't understand the paragraph.*	Surprised by results. *They might be better than they really are, because I just started French two months ago.*	Part I : Context and known words. *I am trying more to compare not by the meaning but by the words themselves.* Part III : Intuition Part IV : Familiar words and logic reasoning.	NC	NC	Translation	Answers *....there you are able to change your response if you think it's not right and here you are not able to do that.*	Clear

BEST COPY AVAILABLE

J.J. (Advanced)	Part I more difficult. *Je sais que Partie IV, c'est plus facile. Partie I est plus long et plus compliqué.*	Surprised : thought she had done better.	Part I : Context. Part II : Personal experience. *Je cherche dans ma tête les phrases que j'ai entendues le plus souvent.* Part III : Try to guess the missing word.	Not long.	Normal	Translation. Personal experience.	Document	Part IV Not clear.
F.R. (Advanced)	Part I more difficult. *...les réponses se ressemblaient et j'étais pas trop sure.*	As expected.	Part III : Reference to L1.	Regular time.	NC	Translation in L1. Grammar (L1). *Je traduis pour me rendre compte si la phrase est bien..., la grammaire, ça ressemble beaucoup à l'espagnol, les verbes,.. on utilise les mêmes en français.*	Document *S'il y a des mots plus difficiles, je dois relire.*	Clear
N.H. (Advanced)	Not difficult. *Not very hard on the brain.*	As expected.	Part I : Context, Intuition, and known words Part III : Intuition *I'm guessing a lot, mostly with the verbs ... I say it loud, to hear what it sounds like.*	Regular time.	Problems with the keyboard.	Translation. Personal experience. *I'm going with that a because I say that a lot.*	Document	Clear
C.L. (Advanced)	Part IV more difficult. *J'ai trouvé la derniere partie la plus difficile, parce que tu peux pas réfléchir.*	As expected.	Part II : Personal experience. Part IV : Try to guess the subject before listening to the audio text.	Not long.	Not at ease. *Je trouve que c'est difficile quand c'est sur ordinateur.*	Translation in L1. Own experience. *C'est comme ça qu'on commence les contes. J'ai travaillé comme fille au pair pendant un mois.*	Document. Answers. *Dans ce cas-ci, sûrement je continue et je retourne à la fin.*	Clear
M.L. (Very advanced)	Too easy.	As expected.	No problem.	Not long.	Problems with the keyboard.	Translation at times.	NC	NC

BEST COPY AVAILABLE